Multiple-Input Multiple-Output Gaussian Broadcast Channels With Confidential Messages

Ruoheng Liu, *Member, IEEE*, Tie Liu, *Member, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

***Abstract*—This paper considers the problem of secret communication over a two-receiver multiple-input multiple-output (MIMO) Gaussian broadcast channel. The transmitter has two independent messages, each of which is intended for one of the receivers but needs to be kept asymptotically perfectly secret from the other. It is shown that, surprisingly, under a matrix power constraint, both messages can be simultaneously transmitted at their respective maximal secrecy rates. To prove this result, the MIMO Gaussian wiretap channel is revisited and a new characterization of its secrecy capacity is provided via a new coding scheme that uses artificial noise (an additive prefix channel) and random binning.**

***Index Terms*—Artificial noise, broadcast channel, channel enhancement, information-theoretic security, multiple-input multiple-output (MIMO) communications, wiretap channel.**

I. Introduction

RAPID advances in wireless technology are quickly moving us toward a pervasively connected world in which a vast array of wireless devices, from iPhones to biosensors, seamlessly communicate with one another. The openness of the wireless medium makes wireless transmission especially susceptible to eavesdropping. Hence, security and privacy issues have become increasingly critical for wireless networks. Although wireless technologies are becoming more and more secure, eavesdroppers are also becoming smarter. Therefore, tackling security at the very basic physical layer is of critical importance.

In this paper, we study the problem of secret communication over a multiple-input multiple-output (MIMO) Gaussian broadcast channel with two receivers. The transmitter is equipped with t transmit antennas, and receiver k, $k = 1, 2$, is equipped with r_k receive antennas. A discrete-time sample of the channel can be written as

$$\mathbf{Y}_k[m] = \mathbf{H}_k \mathbf{X}[m] + \mathbf{Z}_k[m], \quad k = 1, 2 \tag{1}$$

where $\mathbf{H}_k$ is the (real) channel matrix of size $r_k \times t$, and $\{\mathbf{Z}_k[m]\}_m$ is an independent and identically distributed (i.i.d.) additive vector Gaussian noise process with zero mean and identity covariance matrix. The channel input $\{\mathbf{X}[m]\}_m$ is subject to a matrix power constraint

$$\frac{1}{n}\sum_{m=1}^{n} \left(\mathbf{X}[m]\mathbf{X}^{\mathsf{T}}[m]\right) \preceq \mathbf{S} \tag{2}$$

where $\mathbf{S}$ is a positive semidefinite matrix, and "$\preceq$" denotes "less than or equal to" in the positive semidefinite ordering between real symmetric matrices (see Appendix A for some related definitions of semidefinite matrices partial ordering). Note that (2) is a rather general power constraint that subsumes many other important power constraints including the average total and per-antenna power constraints as special cases.

Consider the communication scenario in which there are two independent messages W_1 and W_2 at the transmitter. Message W_1 is intended for receiver 1 but needs to be kept secret from receiver 2, and message W_2 is intended for receiver 1 but needs to be kept secret from receiver 2 (see Fig. 1 for an illustration of this communication scenario). The confidentiality of the messages at the unintended receivers is measured using the normalized information-theoretic criteria (weak secrecy) [1], [2]

$$\frac{1}{n} I(W_1; \mathbf{Y}_2^n) \to 0 \quad \text{and} \quad \frac{1}{n} I(W_2; \mathbf{Y}_1^n) \to 0 \tag{3}$$

where $\mathbf{Y}_k^n := (\mathbf{Y}_k[1], \ldots, \mathbf{Y}_k[n])$, and the limits are taken as the block length $n \to \infty$. The goal is to characterize the entire secrecy rate region $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) = \{(R_1, R_2)\}$ that can be

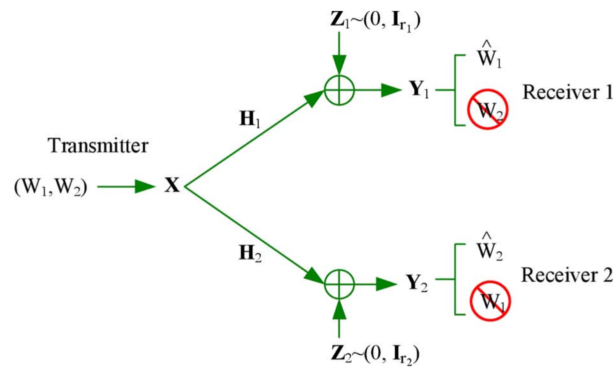

Fig. 1. MIMO Gaussian broadcast channel with confidential messages.

Manuscript received March 23, 2009; revised April 14, 2010. Date of current version August 18, 2010. This work was supported in part by the U.S. National Science Foundation under Grants CCF-07-28208, CCF-08-45848, and CNS-09-05398; in part by the European Commission in the framework of the FP7 Network of Excellence in Wireless Communications NEWCOM++; and in part by the Israel Science Foundation. The material in this paper was presented in part at the 2009 IEEE International Symposium on Information Theory, Seoul, Korea, June 2009.

R. Liu is with Alcatel-Lucent, Murray Hill, NJ 07974 USA (e-mail: ruoheng.liu@alcatel-lucent.com).

T. Liu is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: tieliu@tamu.edu).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

S. Shamai (Shitz) is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: shlomo@ee.technion.ac.il).

Communicated by E. Erkip, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2054593

achieved by any coding scheme. The rate region $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S})$ is usually known as the secrecy capacity region of the channel.

In recent years, information-theoretic study of secret MIMO communication has been an active area of research (see [3] for a recent survey of progress in this area). Most noticeably, the secrecy capacity of the MIMO Gaussian wiretap channel was characterized in [4]–[6] for the multiple-input single-output (MISO) case and [7]–[10] for the general MIMO case. The secrecy capacity region of the MIMO Gaussian broadcast channel with a common and a confidential messages was characterized in [11]. The problem of communicating two confidential messages over the two-receiver MIMO Gaussian broadcast channel was first considered in [12], where it was shown that under the average total power constraint, secret dirty-paper coding (S-DPC) [13] achieves the secrecy capacity region for the MISO case. For the general MIMO case, however, characterizing the secrecy capacity region remained as an open problem.

II. Broadcast Secrecy Channel: Main Results

The main result of this paper is a precise characterization of the secrecy capacity region of the (general) MIMO Gaussian broadcast channel, summarized in the following theorem.

Theorem 1: The secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S})$ of the MIMO Gaussian broadcast channel (1) with confidential messages W_1 (intended for receiver 1 but needing to be kept secret from receiver 2) and W_2 (intended for receiver 2 but needing to be kept secret from receiver 1) under the matrix power constraint (2) is given by the set of nonnegative rate pairs (R_1,R_2) such that

$$R_1\leq \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\intercal}\right|-\frac{1}{2}\log\left|\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\intercal}\right|\right)$$

and

$$R_2\leq \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\frac{\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{S}\mathbf{H}_2^{\intercal}}{\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\intercal}}\right|-\frac{1}{2}\log\left|\frac{\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{S}\mathbf{H}_1^{\intercal}}{\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\intercal}}\right|\right) \tag{4}$$

where $\mathbf{I}_{r_k}$ denotes the identity matrix of size $r_k\times r_k$.

Note that the rate region (4) is *rectangular*. This implies that under the *matrix* power constraint, both confidential messages W_1 and W_2 can be *simultaneously* transmitted at their respective maximal secrecy rates as if over two separate MIMO Gaussian wiretap channels. In other words, the upper bound on R_1 in (4) is the secrecy capacity $C_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S})$[1] of the MIMO Gaussian wiretap channel (1) with receiver 1 being the legitimate receiver and receiver 2 being the eavesdropper; while the upper bound on R_2 in (4) is the secrecy capacity $C_s(\mathbf{H}_2,\mathbf{H}_1,\mathbf{S})$ of the MIMO Gaussian wiretap channel (1) with receiver 2 being the legitimate receiver and receiver 1 being the eavesdropper. Also note that if $\mathbf{B}^\star$ is an optimal solution to the optimization problem

$$\max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\log\left|\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\intercal}\right|-\log\left|\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\intercal}\right|\right) \tag{5}$$

then $\mathbf{B}^\star$ *simultaneously* maximizes both objective functions on the right-hand side (RHS) of (4).

[1]In our notation, the first argument in $C_s(\cdot)$ represents the channel matrix for the legitimate receiver, and the second argument represents the channel matrix for the eavesdropper.

It is rather surprising to see that under the matrix power constraint, both confidential messages W_1 and W_2 can be simultaneously transmitted at their respective maximal secrecy rates over the MIMO Gaussian broadcast channel (1). As we will see, this is due to the fact that there are in fact two efficient coding schemes: one uses only random binning, and the other uses both random binning and *artificial noise*. Both of them can achieve the secrecy capacity of the MIMO Gaussian wiretap channel. Through S-DPC [13], both schemes can be *simultaneously* implemented in communicating confidential messages W_1 and W_2 over the MIMO Gaussian broadcast channel (1).

Remark 1: The secrecy capacity of the MIMO Gaussian wiretap channel under a matrix power constraint was first characterized in [9], by which the secrecy capacity $C_s(\mathbf{H}_2,\mathbf{H}_1,\mathbf{S})$ of the MIMO Gaussian wiretap channel (1) with receiver 2 being the legitimate receiver and receiver 1 being the eavesdropper can also be represented as

$$C_s(\mathbf{H}_2,\mathbf{H}_1,\mathbf{S}) = \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\intercal}\right|-\frac{1}{2}\log\left|\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\intercal}\right|\right). \tag{6}$$

By Theorem 1, the secrecy capacity region of the MIMO Gaussian broadcast channel (1) with confidential messages W_1 and W_2 under the matrix power constraint (2) can also be written as the set of nonnegative rate pairs (R_1,R_2) satisfying

$$R_1\leq \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\intercal}\right|-\frac{1}{2}\log\left|\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\intercal}\right|\right)$$

and

$$R_2\leq \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\intercal}\right|-\frac{1}{2}\log\left|\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\intercal}\right|\right). \tag{7}$$

Note, however, that the optimization problems on the RHS of (7) do not, in general, admit the same optimal solution. As we will see, this makes (4) a better choice when it comes to proving the achievability part of the theorem.

As a corollary of Theorem 1, we have the following characterization of the secrecy capacity region under the average total power constraint. This is a simple consequence of [14, Lemma 1].

Corollary 1: The secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,P)$ of the MIMO Gaussian broadcast channel (1) with confidential messages W_1 (intended for receiver 1 but needing to be kept secret from receiver 2) and W_2 (intended for receiver 2 but needing to be kept secret from receiver 1) under the average total power constraint

$$\frac{1}{n}\sum_{m=1}^{n}\|\mathbf{X}[m]\|^2\leq P \tag{8}$$

is given by

$$\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,P)=\bigcup_{\mathbf{S}\succeq 0,\ \mathrm{Tr}(\mathbf{S})\leq P}\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S}). \tag{9}$$

Remark 2: Unlike Theorem 1, under the average total power constraint, the secrecy capacity region of the MIMO Gaussian broadcast channel is, in general, *not* rectangular. This is because the secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, P)$ is given by the union of $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ over all possible matrix constraints $\mathbf{S}$, and each boundary point of $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, P)$ may correspond to the corner point of $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ for *different* matrix constraints $\mathbf{S}$.

The rest of the paper is devoted to the proof of Theorem 1. As mentioned previously, the rectangular nature of the rate region (4) suggests that the result is intimately connected to the secrecy capacity of the MIMO Gaussian wiretap channel. The secrecy capacity of the MIMO Gaussian wiretap channel under the matrix power constraint was previously characterized in [9], where it was shown that Gaussian random binning *without* prefix coding is optimal. In Section III, we revisit the MIMO Gaussian wiretap channel problem and show that Gaussian random binning *with* prefix coding can also achieve the secrecy capacity, provided that the prefix channel is appropriately chosen. In Section IV, we prove Theorem 1 using two different characterizations of the secrecy capacity of the MIMO Gaussian wiretap channel and S-DPC [13]. Numerical examples are provided in Section V to illustrate the theoretical results. Finally, in Section VI, we conclude the paper with some remarks.

III. MIMO Gaussian Wiretap Channel Revisited

In this section, we revisit the problem of the MIMO Gaussian wiretap channel under a matrix power constraint. The problem was first considered in [9], where a precise characterization of the secrecy capacity was provided. The goal of this section is to provide an alternative characterization of the secrecy capacity which will facilitate the proof of Theorem 1. More specifically, we wish to provide a MIMO wiretap channel bound on the secrecy rate R_2 which will match the RHS of (4).

For that purpose, consider again the MIMO Gaussian broadcast channel (1) but this time with only one confidential message W at the transmitter. Message W is intended for receiver 2 (the legitimate receiver) but needs to be kept secret from receiver 1 (the eavesdropper). The confidentiality of W at receiver 1 is measured using the normalized information-theoretic criteria [1], [2]

$$\frac{1}{n} I(W; \mathbf{Y}_1^n) \to 0. \tag{10}$$

The channel input $\{\mathbf{X}[m]\}_m$ is subject to the matrix power constraint (2). The goal is to characterize the secrecy capacity $C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S})$, which is the maximum achievable secrecy rate for message W. This communication scenario, as illustrated in Fig. 2, is widely known as the MIMO Gaussian wiretap channel [4]–[9].

In their seminal work [2], Csiszár and Körner provided a single-letter characterization of the secrecy capacity

$$C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S}) = \max_{(U,\mathbf{X})} [I(U; \mathbf{Y}_2) - I(U; \mathbf{Y}_1)] \tag{11}$$

where U is an auxiliary variable, and the maximization is over all jointly distributed $(U, \mathbf{X})$ such that $U \to \mathbf{X} \to (\mathbf{Y}_1, \mathbf{Y}_2)$

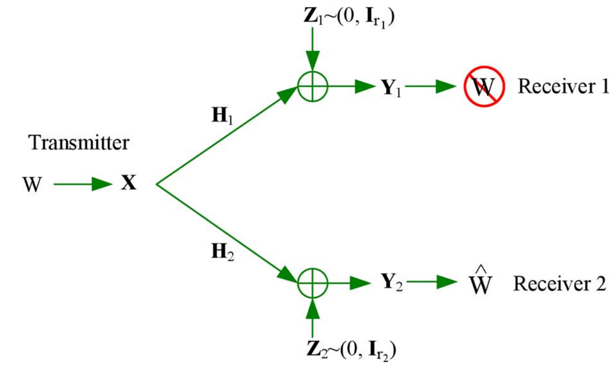

Fig. 2. MIMO Gaussian wiretap channel.

forms a Markov chain and $\mathsf{E}[\mathbf{X}\mathbf{X}^\mathsf{T}] \preceq \mathbf{S}$. Here, $I(U, \mathbf{Y}_k)$ denotes the mutual information between U and $\mathbf{Y}_k$. As shown in [2], the secrecy rate on the RHS of (11) can be achieved by a coding scheme that combines random binning and prefix coding [2]. More specifically, the auxiliary variable U represents a precoding signal, and the conditional distribution of $\mathbf{X}$ given U represents the prefix channel. In [9], Liu and Shamai further studied the optimization problem on the RHS of (11) and showed that a Gaussian $U = \mathbf{X}$ is an optimal solution. Hence, a matrix characterization of the secrecy capacity is given by [9]

$$C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S}) = \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T} \right| - \frac{1}{2} \log \left| \mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T} \right| \right). \tag{12}$$

Thus, Gaussian random binning *without* prefix coding is an optimal coding strategy for the MIMO Gaussian wiretap channel.

Next, we show that a different coding scheme that combines Gaussian random binning *and* prefix coding can also achieve the secrecy capacity of the MIMO Gaussian wiretap channel. This leads to a new characterization of the secrecy capacity as summarized in the following theorem.

Theorem 2: The secrecy capacity $C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S})$ of the MIMO Gaussian broadcast channel (1) with a confidential message W (intended for receiver 2 but needing to be kept secret from receiver 1) under the matrix power constraint (2) is given by

$$C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S}) = \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \frac{\mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{S} \mathbf{H}_2^\mathsf{T}}{\mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T}} \right| - \frac{1}{2} \log \left| \frac{\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{S} \mathbf{H}_1^\mathsf{T}}{\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T}} \right| \right). \tag{13}$$

Remark 3: The achievability of the secrecy rate on the RHS of (13) can be obtained from the Csiszár-Körner expression (11) by choosing $\mathbf{X} = U + V$, where U and V are two independent Gaussian vectors with zero means and covariance matrices $\mathbf{S} - \mathbf{B}$ and $\mathbf{B}$, respectively. This choice of $(U, \mathbf{X})$ differs from that for (12) in two important ways:

1) In (13), the input vector $\mathbf{X}$ always takes the full covariance matrix $\mathbf{S}$. For (12), the covariance matrix of $\mathbf{X}$ needs to be

chosen to solve an optimization program; the full covariance matrix $\mathbf{S}$ is *not* always an optimal solution.

2) In (13), the conditional distribution of $\mathbf{X}$ given U may form a *nontrivial* prefix channel. For (12), $U \equiv \mathbf{X}$ so prefix coding is never applied.

Remark 4: Note that the prefix channel in (13) is an additive vector Gaussian noise channel, so the auxiliary variable V represents an *artificial* noise [15] sent (on purpose) by the transmitter to confuse the eavesdropper. Since the artificial noise has no structure to it, it will add to the noise floor at both legitimate receiver and the eavesdropper.

The converse part of the theorem can be proved using a *channel-enhancement* argument, similar to that in [9]. The details of the proof are provided in Appendix B.

IV. MIMO Gaussian Broadcast Channel With Confidential Messages

In this section, we prove Theorem 1. To prove the converse part of the theorem, we will consider a single-message, wiretap channel bound on the secrecy rates R_1 and R_2. More specifically, note that both messages W_1 and W_2 can be transmitted at the maximum secrecy rate when the other message is absent from the transmission. Therefore, to bound from above the secrecy rate R_1, we assume that only W_1 needs to be communicated over the channel. This is precisely a MIMO Gaussian wiretap channel problem with receiver 1 as legitimate receiver and receiver 2 as eavesdropper. Reversing the roles of receiver 1 and 2, we have from (12) that

$$\begin{aligned} R_1 &\le C_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) \\ &= \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T} \right| - \frac{1}{2} \log \left| \mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T} \right| \right). \end{aligned} \tag{14}$$

Similarly, to bound from above the secrecy rate R_2, let us assume that only W_2 needs to be communicated over the channel. This is, again, a MIMO Gaussian wiretap channel problem with receiver 2 playing the role of legitimate receiver and receiver 1 playing the role of eavesdropper. By Theorem 2

$$\begin{aligned} R_2 &\le C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S}) \\ &= \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \frac{\mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{S} \mathbf{H}_2^\mathsf{T}}{\mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T}} \right| - \frac{1}{2} \log \left| \frac{\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{S} \mathbf{H}_1^\mathsf{T}}{\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T}} \right| \right). \end{aligned} \tag{15}$$

Putting together (14) and (15), we have proved the converse part of the theorem.

Next, we show that every rate pair (R_1, R_2) within the secrecy rate region (4) is achievable. Note that (4) is rectangular, so we only need to show that the corner point (R_1, R_2) given by

$$R_1 = \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T} \right| - \frac{1}{2} \log \left| \mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T} \right| \right)$$

and

$$R_2 = \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \frac{\mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{S} \mathbf{H}_2^\mathsf{T}}{\mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T}} \right| - \frac{1}{2} \log \left| \frac{\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{S} \mathbf{H}_1^\mathsf{T}}{\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T}} \right| \right) \tag{16}$$

is achievable.

Recall from [13] that for any jointly distributed $(V_1, V_2, \mathbf{X})$ such that $(V_1, V_2) \to \mathbf{X} \to (\mathbf{Y}_1, \mathbf{Y}_2)$ forms a Markov chain and $\mathsf{E}[\mathbf{X}\mathbf{X}^\mathsf{T}] \preceq \mathbf{S}$, the secrecy rate pair (R_1, R_2) given by

$$\begin{aligned} R_1 &= I(V_1; \mathbf{Y}_1) - I(V_1; V_2, \mathbf{Y}_2) \\ \text{and} \quad R_2 &= I(V_2; \mathbf{Y}_2) - I(V_2; V_1, \mathbf{Y}_1) \end{aligned} \tag{17}$$

is achievable for the MIMO Gaussian broadcast channel (1) under the matrix power constraint (2). In [13], the achievability of the rate pair (17) was proved using a *double-binning* scheme. Specifically, the auxiliary variables V_1 and V_2 represent the precoding signals for the confidential messages W_1 and W_2, respectively.

Now let $\mathbf{B}$ be a positive semidefinite matrix such that $\mathbf{B} \preceq \mathbf{S}$, and let

$$\begin{aligned} V_1 &= \mathbf{U}_1 + \mathbf{F}\mathbf{U}_2 \\ V_2 &= \mathbf{U}_2 \\ \text{and} \quad \mathbf{X} &= \mathbf{U}_1 + \mathbf{U}_2 \end{aligned} \tag{18}$$

where $\mathbf{U}_1$ and $\mathbf{U}_2$ are two independent Gaussian vectors with zero means and covariance matrices $\mathbf{B}$ and $\mathbf{S} - \mathbf{B}$, respectively, and

$$\mathbf{F} := \mathbf{B}\mathbf{H}_1^\mathsf{T}(\mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T})^{-1} \mathbf{H}_1. \tag{19}$$

By (18)

$$\mathbf{Y}_k = \mathbf{H}_k(\mathbf{U}_1 + \mathbf{U}_2) + \mathbf{Z}_k \tag{20}$$

for $k = 1, 2$. Note that the matrix $\mathbf{F}$ defined in (19) is precisely the *precoding* matrix for suppressing $\mathbf{U}_2$ from $\mathbf{Y}_1$ [16, Theorem 1]. Hence

$$\begin{aligned} I(V_1; \mathbf{Y}_1) - I(V_1; V_2) &= I(V_1; \mathbf{Y}_1) - I(V_1; \mathbf{U}_2) \\ &= \frac{1}{2} \log \left| \mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T} \right|. \end{aligned} \tag{21}$$

Moreover

$$\begin{aligned} I(V_1; \mathbf{Y}_2 | V_2) &= I(\mathbf{U}_1 + \mathbf{F}\mathbf{U}_2; \mathbf{H}_2(\mathbf{U}_1 + \mathbf{U}_2) + \mathbf{Z}_2 | \mathbf{U}_2) \\ &= I(\mathbf{U}_1; \mathbf{H}_2 \mathbf{U}_1 + \mathbf{Z}_2 | \mathbf{U}_2) \\ &= I(\mathbf{U}_1; \mathbf{H}_2 \mathbf{U}_1 + \mathbf{Z}_2) \\ &= \frac{1}{2} \log \left| \mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T} \right| \end{aligned} \tag{22}$$

where the third equality follows from the fact that $\mathbf{U}_1$ and $\mathbf{U}_2$ are independent. Putting together (21) and (22), we have

$$\begin{aligned} &I(V_1; \mathbf{Y}_1) - I(V_1; V_2, \mathbf{Y}_2) \\ &\quad = [I(V_1; \mathbf{Y}_1) - I(V_1; V_2)] - I(V_1; \mathbf{Y}_2 | V_2) \\ &\quad = \frac{1}{2} \log \left| \mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{B} \mathbf{H}_1^\mathsf{T} \right| - \frac{1}{2} \log \left| \mathbf{I}_{r_2} + \mathbf{H}_2 \mathbf{B} \mathbf{H}_2^\mathsf{T} \right|. \end{aligned} \tag{23}$$

Similarly

$$\begin{aligned} I(V_1, V_2; \mathbf{Y}_1) &= I(\mathbf{U}_1 + \mathbf{F}\mathbf{U}_2, \mathbf{U}_2; \mathbf{H}_1(\mathbf{U}_1 + \mathbf{U}_2) + \mathbf{Z}_2) \\ &= I(\mathbf{U}_1, \mathbf{U}_2; \mathbf{H}_1(\mathbf{U}_1 + \mathbf{U}_2) + \mathbf{Z}_2) \\ &= \frac{1}{2} \log \left| \mathbf{I}_{r_1} + \mathbf{H}_1 \mathbf{S} \mathbf{H}_1^\mathsf{T} \right|. \end{aligned} \tag{24}$$

Thus

$$\begin{aligned} I(V_2;V_1,\mathbf{Y}_1) &= I(V_2;\mathbf{Y}_1|V_1) + I(V_2;V_1) \\ &= I(V_1,V_2;\mathbf{Y}_1) - [I(V_1;\mathbf{Y}_1) - I(V_1;V_2)] \\ &= \frac{1}{2}\log\left|\frac{\mathbf{I}_{r_1} + \mathbf{H}_1\mathbf{S}\mathbf{H}_1^{\mathsf{T}}}{\mathbf{I}_{r_1} + \mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\mathsf{T}}}\right| \end{aligned} \tag{25}$$

where the last equality follows from (21) and (24). Moreover

$$\begin{aligned} I(V_2;\mathbf{Y}_2) &= I(\mathbf{U}_2;\mathbf{H}_2(\mathbf{U}_1+\mathbf{U}_2)+\mathbf{Z}_2) \\ &= \frac{1}{2}\log\left|\frac{\mathbf{I}_{r_2} + \mathbf{H}_2\mathbf{S}\mathbf{H}_2^{\mathsf{T}}}{\mathbf{I}_{r_2} + \mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\mathsf{T}}}\right|. \end{aligned} \tag{26}$$

Putting together (25) and (26), we have

$$\begin{aligned} &I(V_2;\mathbf{Y}_2) - I(V_2;V_1,\mathbf{Y}_1) \\ &= \frac{1}{2}\log\left|\frac{\mathbf{I}_{r_2} + \mathbf{H}_2\mathbf{S}\mathbf{H}_2^{\mathsf{T}}}{\mathbf{I}_{r_2} + \mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\mathsf{T}}}\right| - \frac{1}{2}\log\left|\frac{\mathbf{I}_{r_1} + \mathbf{H}_1\mathbf{S}\mathbf{H}_1^{\mathsf{T}}}{\mathbf{I}_{r_1} + \mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\mathsf{T}}}\right|. \end{aligned} \tag{27}$$

Finally, let $\mathbf{B}$ be an optimal solution to the optimization program (5). As mentioned previously in Section II, such a choice will *simultaneously* maximize the RHS of (23) and (27). Thus, the corner point (16) is indeed achievable. This completes the proof of the theorem.

Remark 5: Note that in standard dirty-paper coding (DPC), the precoding matrix $\mathbf{F}$ is chosen to cancel the known interference. In our scheme, such a choice plays two important roles. First, it helps to cancel the precoding signal representing message W_2, so message W_1 sees an interference-free legitimate receiver channel. Second, it helps to boost the security for message W_2 by causing interference to the corresponding eavesdropper. For this reason, we call our scheme S-DPC, to differentiate it from the standard DPC.

Remark 6: In S-DPC, both the legitimate receiver and the eavesdropper for message W_1 are interference free. On the other hand, for message W_2, both the legitimate receiver and the eavesdropper are subject to interference from the precoding signal representing message W_1. As we have seen in Section III, the secrecy capacity of the MIMO Gaussian wiretap channel can be achieved with or without interference in place. Therefore, both secrecy capacity achieving schemes can be simultaneously implemented via S-DPC to simultaneously communicate both confidential messages at their respective maximal secrecy rates.

V. Computation of Secrecy Capacity and Numerical Examples

In this section, we provide numerical examples to illustrate the secrecy capacity region of the MIMO Gaussian wiretap channel with confidential messages. As shown in (4) and (9), under both matrix and average total power constraints, the secrecy capacity regions $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S})$ and $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,P)$ are expressed in terms of matrix optimization programs (though implicit in (9)). In general, these optimization programs are not convex, and hence, finding the boundary of the secrecy capacity regions is nontrivial.

In [12], a precise characterization of the secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,P)$ was obtained for the *MISO* Gaussian broadcast channel using the generalized eigenvalue decomposition [17, Ch. 6.3]. For the *aligned* MIMO Gaussian wiretap channel, [10] provided an explicit, closed-form expression for the secrecy capacity. In the following, we generalize the results of [10] and [12] to the general MIMO Gaussian broadcast channel under the matrix power constraint.

Let ϕ_j, $j=1,\dots,t$, be the generalized eigenvalues of the pencil (see Appendix A for the definition of matrix pencil)

$$\left(\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^{\mathsf{T}}\mathbf{H}_1\mathbf{S}^{\frac{1}{2}},\ \mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{H}_2^{\mathsf{T}}\mathbf{H}_2\mathbf{S}^{\frac{1}{2}}\right). \tag{28}$$

Since both $\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{H}_1^{\mathsf{T}}\mathbf{H}_1\mathbf{S}^{\frac{1}{2}}$ and $\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{H}_2^{\mathsf{T}}\mathbf{H}_2\mathbf{S}^{\frac{1}{2}}$ are strictly positive definite, we have $\phi_j > 0$ for $j=1,\dots,t$. Without loss of generality, we may assume that these generalized eigenvalues are ordered as

$$\phi_1 \ge \cdots \ge \phi_\rho > 1 \ge \phi_{\rho+1} \ge \cdots \ge \phi_t > 0 \tag{29}$$

i.e., a total of ρ of them are assumed to be greater than 1. We have the following characterization of the secrecy capacity of the MIMO Gaussian wiretap channel under the matrix power constraint, which is a natural extension of [10].

Theorem 3: The secrecy capacity $C_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S})$ of the MIMO Gaussian broadcast channel (1) with confidential message W (intended for receiver 1 but needing to be kept secret from receiver 2) under the matrix power constraint (2) is given by

$$C_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S}) = \frac{1}{2}\sum_{j=1}^{\rho}\log\phi_j \tag{30}$$

where ϕ_j, $j=1,\dots,\rho$, are the generalized eigenvalues of the pencil (28) that are greater than 1.

Remark 7: Note that $\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{H}_2^{\mathsf{T}}\mathbf{H}_2\mathbf{S}^{\frac{1}{2}}$ is invertible, so computing the generalized eigenvalues of the pencil (28) can be reduced to the problem of finding standard eigenvalues of a related semidefinite matrix [17, Ch. 6.3]. Hence, the secrecy capacity expression (30) is computable.

A proof of the theorem following the approach of [10] is provided in Appendix C. As a corollary, we have the following characterization of the secrecy capacity region of the MIMO Gaussian broadcast channel with confidential messages under the matrix power constraint.

Corollary 2: The secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1,\mathbf{H}_2,\mathbf{S})$ of the MIMO Gaussian broadcast channel (1) with confidential messages W_1 (intended for receiver 1 but needing to be kept secret from receiver 2) and W_2 (intended for receiver 2 but needing to be kept secret from receiver 1) under the matrix constraint (2) is given by the set of nonnegative rate pairs (R_1,R_2) such that

$$\begin{aligned} R_1 &\le \frac{1}{2}\sum_{j=1}^{\rho}\log\phi_j \\ \text{and}\quad R_2 &\le \frac{1}{2}\sum_{j=\rho+1}^{t}\log\frac{1}{\phi_j} \end{aligned} \tag{31}$$

where ϕ_j, $j=1,\dots,\rho$, are the generalized eigenvalues of the pencil (28) that are greater than 1, and ϕ_j, $j=\rho+1,\dots,t$, are the generalized eigenvalues of the pencil (28) that are less than or equal to 1.

A proof of Corollary 2 is deferred to Appendix D. Under the average total power constraint, we have not been able to find a computable secrecy capacity expression for the general MIMO case. We can, however, write the secrecy capacity region under the average total power constraint as in (9). For any given

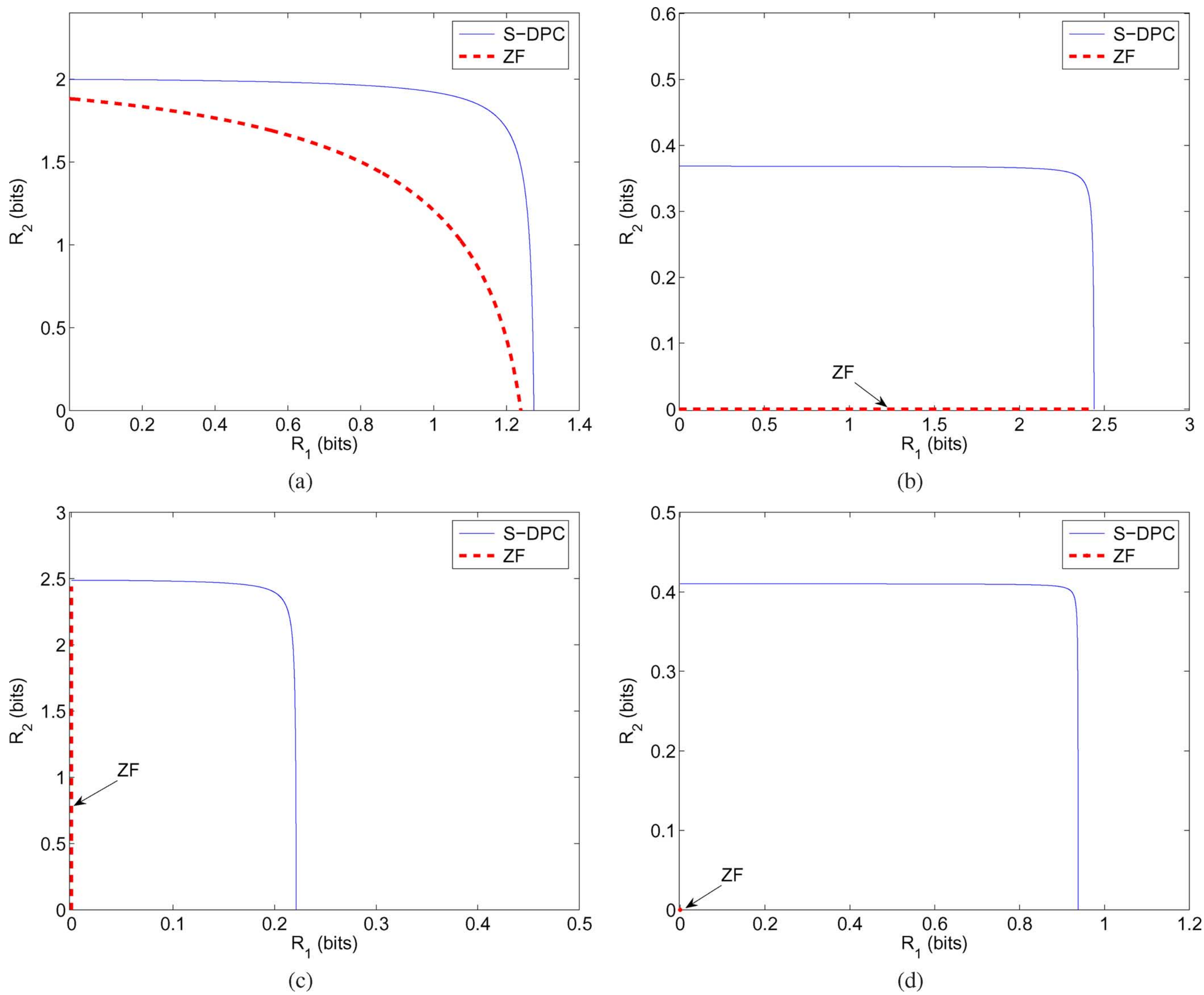

Fig. 3. Secrecy rate regions of the MIMO Gaussian broadcast channel under the average total power constraint. (a) $C_s(\mathbf{h}_{11}, \mathbf{h}_{22}, P)$ ($r_1 = r_2 = 1$). (b)$C_s(\mathbf{H}_1, \mathbf{h}_{22}, P)$ ($r_1 = 2, r_2 = 1$). (c) $C_s(\mathbf{h}_{11}, \mathbf{H}_2, P)$ ($r_1 = 1, r_2 = 2$). (d) $C_s(\mathbf{H}_1, \mathbf{H}_2, P)$ ($r_1 = r_2 = 2$).

matrix constraint $\mathbf{S}$, $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ can be computed as given by (31). Then, the secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, P)$ can be found through an exhaustive search over the set of matrices $\{\mathbf{S} : \mathbf{S} \succeq 0 \text{ and } \mathsf{Tr}(\mathbf{S}) \le P\}$. Note that, each boundary point $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, P)$ may correspond to the corner point of $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ for different matrix constraints S. Hence, under the average total power constraint, the secrecy capacity region of the MIMO Gaussian broadcast channel is not rectangular in general.

Let $\mathbf{h}_{11} = (0.3\ 2.5)$, $\mathbf{h}_{12} = (2.2\ 1.8)$, $\mathbf{h}_{21} = (1.3\ 1.2)$, $\mathbf{h}_{22} = (1.5\ 3.9)$ and $P = 12$, and let

$$\mathbf{H}_k = \begin{pmatrix} \mathbf{h}_{k1} \\ \mathbf{h}_{k2} \end{pmatrix}, \quad k = 1, 2. \tag{32}$$

The secrecy capacity regions $\mathcal{C}_s(\mathbf{h}_{11}, \mathbf{h}_{22}, P)$, $\mathcal{C}_s(\mathbf{H}_1, \mathbf{h}_{22}, P)$, $\mathcal{C}_s(\mathbf{h}_{11}, \mathbf{H}_2, P)$ and $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, P)$ are illustrated in Fig. 3. For comparison, we have also plotted the secrecy rate regions achieved by the simple zero-forcing (ZF) strategy. In ZF, each of the confidential messages is encoded using a vector Gaussian signal. To guarantee confidentiality, the covariance matrices of the transmit signals are chosen in the *null* space of the channel matrix at the unintended receiver. Hence, the achievable secrecy rate region is given by (33), shown at the bottom of the page. Note that unlike the secrecy capacity region expression (9), computing the rate region (33) only involves solving convex optimization programs. As shown in Fig. 3, in all four scenarios, ZF is strictly suboptimal as compared with

$$\mathcal{R}_S^{\mathrm{ZF}}(\mathbf{H}_1, \mathbf{H}_2, P) = \bigcup_{\substack{\mathbf{B}_1 \succeq 0,\ \mathbf{B}_2 \succeq 0,\ \mathsf{Tr}(\mathbf{B}_1 + \mathbf{B}_2) \le P \\ \mathbf{H}_2\mathbf{B}_1 = 0,\ \mathbf{H}_1\mathbf{B}_2 = 0}} \left\{ (R_1, R_2) \middle| \begin{array}{l} R_1 \le \frac{1}{2}\log|\mathbf{I}_{r_1} + \mathbf{H}_1\mathbf{B}_1\mathbf{H}_1^{\mathsf{T}}| \\ R_2 \le \frac{1}{2}\log|\mathbf{I}_{r_2} + \mathbf{H}_2\mathbf{B}_2\mathbf{H}_2^{\mathsf{T}}| \end{array} \right\} \tag{33}$$

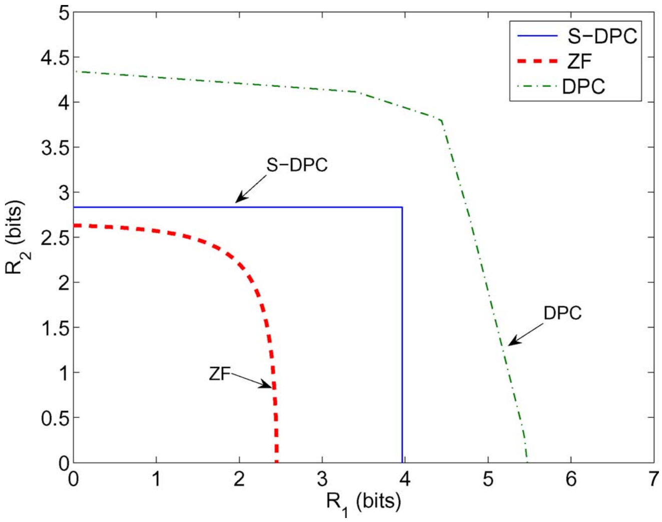

Fig. 4. Rate regions of the MIMO Gaussian broadcast channel under the power matrix constraint.

S-DPC. In particular, if the channel matrix of the unintended receiver has full row rank, ZF cannot achieve any positive secrecy rate for the corresponding confidential message. On the other hand, S-DPC can always achieve positive secrecy rates for both confidential messages unless the MIMO Gaussian broadcast channel is degraded.

Finally, let

$$\mathbf{H}_1 = \begin{pmatrix} 1.8 & -2.0 & 2.0 \\ 1.0 & -6.0 & 3.0 \end{pmatrix} \tag{34}$$

$$\mathbf{H}_2 = \begin{pmatrix} 2.3 & 2.0 & -3 \\ 2.0 & 1.2 & -1.5 \end{pmatrix} \tag{35}$$

and

$$\mathbf{S} = \begin{pmatrix} 5.0 & -0.7 & -2.0 \\ -0.7 & 3.8 & -2.5 \\ -2.0 & -2.5 & 5.0 \end{pmatrix}. \tag{36}$$

Fig. 4 illustrates the secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ of the MIMO Gaussian broadcast channel (1) under the matrix power constraint (2). Here, the secrecy capacity region $\mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ is plotted based on the computable expression (31). Also in the figure are the secrecy rate region $\mathcal{R}_s^{\mathrm{ZF}}(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ achieved by ZF strategy and the *nonsecrecy* capacity region $\mathcal{C}(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$ achieved by standard DPC [14]. As expected, we have $\mathcal{R}_s^{\mathrm{ZF}}(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) \subset \mathcal{C}_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) \subset \mathcal{C}(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S})$.

VI. Concluding Remarks

In this paper, we have considered the problem of communicating two confidential messages over a two-receiver MIMO Gaussian broadcast channel. Each of the confidential messages is intended for one of the receivers but needs to be kept asymptotically perfectly secret from the other. Precise characterizations of the secrecy capacity region have been provided under both matrix and average total power constraints. Surprisingly, under the matrix power constraint, both confidential messages can be transmitted simultaneously at their respective maximal secrecy rates.

To prove this result, we have revisited the problem of the MIMO Gaussian wiretap channel and proposed a new coding scheme that achieves the secrecy capacity of the channel. Unlike the previous scheme considered in [4]–[9] where prefix coding is not applied, the new coding scheme uses artificial vector Gaussian noise as a way of prefix coding. Moreover, the optimal covariance matrix of the artificial noise coincides with that of the transmit signal in the previous scheme (with a reversal of the roles of legitimate receiver and eavesdropper). This allows both schemes to be overlayed via secret dirty-paper coding without sacrificing the secrecy rate performance for either of them. We believe that the new insights into the MIMO Gaussian wiretap channel problem gained in this work will help to solve some other MIMO Gaussian secret communication problems.

Appendix A
Some Matrix Notation and Definitions

Below we summarize various properties of symmetric positive semidefinite matrices that are used in this paper.

1) Semidefinite Matrices and Partial Ordering: An $n \times n$ real symmetric matrix $\mathbf{A}$ is positive semidefinite if $\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} \geq 0$ for all nonzero vectors $\mathbf{x}$ with real entries ($\mathbf{x} \in \mathbb{R}^n$), where $\mathbf{x}^{\mathsf{T}}$ denotes the transpose of $\mathbf{x}$; $\mathbf{A}$ is positive definite if $\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} > 0$ for all nonzero vectors $\mathbf{x}$ with real entries.

If $\mathbf{A}$ is positive semidefinite, we write $\mathbf{A} \succeq 0$ and if $\mathbf{A}$ is positive-definite we writes $\mathbf{A} \succ 0$. For arbitrary square $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$, we write $\mathbf{A} \succeq \mathbf{B}$ (or $\mathbf{B} \preceq \mathbf{A}$) if $\mathbf{A} - \mathbf{B} \succeq 0$, i.e., $\mathbf{A} - \mathbf{B}$ is positive semidefinite. This defines a partial ordering on the set of all real symmetric matrices. Similarly, we define a strict partial ordering $\mathbf{A} \succ \mathbf{B}$ (or $\mathbf{B} \prec \mathbf{A}$) if if $\mathbf{A} - \mathbf{B} \succ 0$.

We have the following properties of positive semidefinite matrices:

1) Every positive definite matrix is invertible, and its inverse is also positive definite. Furthermore, if $\mathbf{A} \succeq \mathbf{B} \succ 0$ then $\mathbf{B}^{-1} \succeq \mathbf{A}^{-1} \succ 0$.
2) If $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite, then the sum $\mathbf{A} + \mathbf{B}$ and the products $\mathbf{ABA}$ and $\mathbf{BAB}$ are also positive semidefinite. If $\mathbf{AB} = \mathbf{BA}$, then $\mathbf{AB}$ is also positive semidefinite.
3) A matrix $\mathbf{A}$ is positive semidefinite if and only if there is a positive semidefinite matrix $\mathbf{M}$ with $\mathbf{M}^2 = \mathbf{A}$. This matrix $\mathbf{M}$, called the square root of $\mathbf{A}$, is unique and is denoted by $\mathbf{M} = \mathbf{A}^{1/2}$. If $\mathbf{A} \succeq \mathbf{B} \succeq 0$ then $\mathbf{A}^{1/2} \succeq \mathbf{B}^{1/2} \succeq 0$.

2) Matrix Pencil and Generalized Eigenvalue: A linear matrix pencil is a matrix-valued function of the form $\mathbf{A} - \lambda\mathbf{B}$ with $\lambda \in \mathbb{R}$, where $\mathbf{A}$ and $\mathbf{B}$ are real $n \times n$ matrices, and is denoted by $(\mathbf{A}, \mathbf{B})$. The eigenvalues of a pencil $(\mathbf{A}, \mathbf{B})$ consist of all real numbers λ for which $|\mathbf{A} - \lambda\mathbf{B}| = 0$. The problem of determining the eigenvalues of a matrix pencil $(\mathbf{A}, \mathbf{B})$ can be equivalently formulated as the generalized eigenvalue problem of determining the nontrivial solutions of the equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}. \tag{37}$$

The nonzero values of λ that satisfy (37) are the generalized eigenvalues and the corresponding vectors $\mathbf{x}$ are the generalized eigenvectors. In particular, if $\mathbf{A}$ is symmetric and $\mathbf{B}$ is symmetric and positive definite, the generalized eigenvalues are all real.

APPENDIX B
PROOF OF THEOREM 2

In this Appendix, we prove Theorem 2. As mentioned previously in Remark 3, the secrecy rate on the RHS of (13) can be achieved by a coding scheme that combines Gaussian random binning and prefix coding. We, therefore, concentrate on proving the converse part of the theorem.

Following [9], we will first prove the converse result for the special case where the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are square and invertible. Next, we will broaden the result to the general case by approximating arbitrary channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ by square and invertible ones. For brevity, we will term the special case as the aligned MIMO Gaussian wiretap channel and the general case as the general MIMO Gaussian wiretap channel.

1) Aligned MIMO Gaussian Wiretap Channel: Consider the special case of the MIMO Gaussian broadcast channel (1) where the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are square and invertible. Multiplying both sides of (1) by $\mathbf{H}_k^{-1}$, the channel model can be equivalently written as

$$\mathbf{Y}_k[m] = \mathbf{X}[m] + \mathbf{Z}_k[m], \quad k = 1, 2 \tag{38}$$

where $\{\mathbf{Z}_k[m]\}_m$ is an i.i.d. additive vector Gaussian noise process with zero mean and covariance matrix

$$\mathbf{N}_k = \mathbf{H}_k^{-1}\mathbf{H}_k^{-\mathsf{T}}. \tag{39}$$

Denote by $C_s(\mathbf{N}_2, \mathbf{N}_1, \mathbf{S})$ the secrecy capacity of (38) (viewed as a MIMO Gaussian wiretap channel with receiver 2 as legitimate receiver and receiver 1 as eavesdropper) under the matrix power constraint (2). We have the following characterization of $C_s(\mathbf{N}_2, \mathbf{N}_1, \mathbf{S})$.

Lemma 1: The secrecy capacity

$$C_s(\mathbf{N}_2, \mathbf{N}_1, \mathbf{S}) = \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \frac{\mathbf{S}+\mathbf{N}_2}{\mathbf{B}+\mathbf{N}_2} \right| - \frac{1}{2} \log \left| \frac{\mathbf{S}+\mathbf{N}_1}{\mathbf{B}+\mathbf{N}_1} \right| \right). \tag{40}$$

Proof: The achievability of the secrecy rate on the RHS of (40) can be obtained from the Csiszár-Körner expression (11) by choosing $\mathbf{X} = U + V$, where U and V are two independent Gaussian vectors with zero means and covariance matrices $\mathbf{S} - \mathbf{B}$ and $\mathbf{B}$, respectively.

To prove the converse result, we will follow [9] and consider a channel-enhancement argument [14] as follows. Let us first assume that $\mathbf{S} \succ 0$. In this case, let $\mathbf{B}^\star$ be an optimal solution to the optimization problem on the RHS of (40). Then, $\mathbf{B}^\star$ must satisfy the following Karush–Kuhn–Tucker conditions:

$$(\mathbf{B}^\star + \mathbf{N}_1)^{-1} + \mathbf{M}_1 = (\mathbf{B}^\star + \mathbf{N}_2)^{-1} + \mathbf{M}_2 \tag{41a}$$
$$\mathbf{B}^\star \mathbf{M}_1 = 0 \tag{41b}$$
$$\text{and} \quad (\mathbf{S} - \mathbf{B}^\star)\mathbf{M}_2 = 0 \tag{41c}$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are positive semidefinite matrices. Let $\widetilde{\mathbf{N}}_1$ be a real symmetric matrix such that

$$\widetilde{\mathbf{N}}_1 := \left(\mathbf{N}_1^{-1} + \mathbf{M}_1\right)^{-1}. \tag{42}$$

Note that $\mathbf{N}_1 \succ 0$ and $\mathbf{M}_1 \succeq 0$. By the above definition of $\widetilde{\mathbf{N}}_1$, we have

$$0 \prec \widetilde{\mathbf{N}}_1 \preceq \mathbf{N}_1 \quad . \tag{43}$$

Since $\mathbf{B}^\star \mathbf{M}_1 = 0$, we have $\mathbf{M}_1 \mathbf{B}^\star = 0$ and thus

$$\begin{aligned}
&[(\mathbf{B}^\star + \mathbf{N}_1)^{-1} + \mathbf{M}_1]^{-1} \\
&= (\mathbf{I}_t + \mathbf{N}_1\mathbf{M}_1)^{-1} (\mathbf{B}^\star + \mathbf{N}_1) \\
&= (\mathbf{I}_t + \mathbf{N}_1\mathbf{M}_1)^{-1} (\mathbf{B}^\star + \mathbf{N}_1) - \mathbf{B}^\star + \mathbf{B}^\star &(44)\\
&= (\mathbf{I}_t + \mathbf{N}_1\mathbf{M}_1)^{-1} [(\mathbf{B}^\star + \mathbf{N}_1) - (\mathbf{I}_t + \mathbf{N}_1\mathbf{M}_1)\mathbf{B}^\star] + \mathbf{B}^\star \\
&= (\mathbf{I}_t + \mathbf{N}_1\mathbf{M}_1)^{-1} \mathbf{N}_1 + \mathbf{B}^\star \\
&= \left(\mathbf{N}_1^{-1} + \mathbf{M}_1\right)^{-1} + \mathbf{B}^\star \\
&= \widetilde{\mathbf{N}}_1 + \mathbf{B}^\star &(45)
\end{aligned}$$

where the last step follows from the definition of $\widetilde{\mathbf{N}}_1$ in (42). Therefore

$$(\mathbf{B}^\star + \widetilde{\mathbf{N}}_1)^{-1} = (\mathbf{B}^\star + \mathbf{N}_1)^{-1} + \mathbf{M}_1. \tag{46}$$

By the KKT condition (41a), we have

$$(\mathbf{B}^\star + \widetilde{\mathbf{N}}_1)^{-1} = (\mathbf{B}^\star + \mathbf{N}_2)^{-1} + \mathbf{M}_2. \tag{47}$$

which implies

$$0 \prec \widetilde{\mathbf{N}}_1 \preceq \mathbf{N}_2. \tag{48}$$

Since $\mathbf{B}^\star \mathbf{M}_1 = 0$, (left) multiplying both sides of (46) by $\mathbf{B}^\star$, we obtain

$$\widetilde{\mathbf{N}}_1(\mathbf{B}^\star + \widetilde{\mathbf{N}}_1)^{-1} = \mathbf{N}_1(\mathbf{B}^\star + \mathbf{N}_1)^{-1} \tag{49}$$

and thus

$$\left| \frac{\mathbf{B}^\star + \widetilde{\mathbf{N}}_1}{\widetilde{\mathbf{N}}_1} \right| = \left| \frac{\mathbf{B}^\star + \mathbf{N}_1}{\mathbf{N}_1} \right|. \tag{50}$$

Similarly, since $(\mathbf{S} - \mathbf{B}^\star)\mathbf{M}_2 = 0$, (left) multiplying both sides of (46) by $\mathbf{S} - \mathbf{B}^\star$, we obtain

$$(\mathbf{S} + \widetilde{\mathbf{N}}_1)(\mathbf{B}^\star + \widetilde{\mathbf{N}}_1)^{-1} = (\mathbf{S} + \mathbf{N}_2)(\mathbf{B}^\star + \mathbf{N}_2)^{-1} \tag{51}$$

and hence

$$\left| \frac{\mathbf{S} + \widetilde{\mathbf{N}}_1}{\mathbf{B}^\star + \widetilde{\mathbf{N}}_1} \right| = \left| \frac{\mathbf{S} + \mathbf{N}_2}{\mathbf{B}^\star + \mathbf{N}_2} \right|. \tag{52}$$

Now consider an enhanced MIMO Gaussian broadcast channel

$$\mathbf{Y}_1[m] = \mathbf{X}[m] + \mathbf{Z}_1[m]$$
$$\text{and} \quad \mathbf{Y}_2[m] = \mathbf{X}[m] + \tilde{\mathbf{Z}}_1[m] \tag{53}$$

where $\{\mathbf{Z}_1[m]\}_m$ and $\{\tilde{\mathbf{Z}}_1[m]\}_m$ are i.i.d. additive vector Gaussian noise processes with zero means and covariance matrices $\mathbf{N}_1$ and $\widetilde{\mathbf{N}}_1$, respectively. Denote by $C_s(\widetilde{\mathbf{N}}_1, \mathbf{N}_1, \mathbf{S})$ the secrecy capacity of (53) (viewed as a MIMO Gaussian wiretap channel with receiver 2 as legitimate receiver and receiver 1 as eavesdropper) under the matrix constraint (2). Note from (43) that $\widetilde{\mathbf{N}}_1 \preceq \mathbf{N}_1$, so the enhanced MIMO Gaussian wiretap channel (53) is *degraded*. Hence

$$\begin{aligned} C_s(\widetilde{\mathbf{N}}_1, \mathbf{N}_1, \mathbf{S}) &= \frac{1}{2}\log\left|\frac{\mathbf{S}+\widetilde{\mathbf{N}}_1}{\widetilde{\mathbf{N}}_1}\right| - \frac{1}{2}\log\left|\frac{\mathbf{S}+\mathbf{N}_1}{\mathbf{N}_1}\right| \\ &= \frac{1}{2}\log\left(\left|\frac{\mathbf{S}+\widetilde{\mathbf{N}}_1}{\mathbf{S}+\mathbf{N}_1}\right|\left|\frac{\mathbf{N}_1}{\widetilde{\mathbf{N}}_1}\right|\right) \\ &= \frac{1}{2}\log\left(\left|\frac{\mathbf{S}+\widetilde{\mathbf{N}}_1}{\mathbf{S}+\mathbf{N}_1}\right|\left|\frac{\mathbf{B}^\star+\mathbf{N}_1}{\mathbf{B}^\star+\widetilde{\mathbf{N}}_1}\right|\right) \\ &= \frac{1}{2}\log\left(\left|\frac{\mathbf{S}+\widetilde{\mathbf{N}}_1}{\mathbf{B}^\star+\widetilde{\mathbf{N}}_1}\right|\left|\frac{\mathbf{B}^\star+\mathbf{N}_1}{\mathbf{S}+\mathbf{N}_1}\right|\right) \\ &= \frac{1}{2}\log\left(\left|\frac{\mathbf{S}+\mathbf{N}_2}{\mathbf{B}^\star+\mathbf{N}_2}\right|\left|\frac{\mathbf{B}^\star+\mathbf{N}_1}{\mathbf{S}+\mathbf{N}_1}\right|\right) \\ &= \frac{1}{2}\log\left|\frac{\mathbf{S}+\mathbf{N}_2}{\mathbf{B}^\star+\mathbf{N}_2}\right| - \frac{1}{2}\log\left|\frac{\mathbf{S}+\mathbf{N}_1}{\mathbf{B}^\star+\mathbf{N}_1}\right| \end{aligned} \tag{54}$$

where the first equality follows from [9, Theorem 1]; the third equality follows from (50); and the fifth equality follows from (52).

Finally, note from (48) that $\widetilde{\mathbf{N}}_1 \preceq \mathbf{N}_2$, i.e., the legitimate receiver in the enhanced wiretap channel (53) receives a better signal that the legitimate receiver in the original wiretap channel (38). Therefore

$$C_s(\mathbf{N}_2, \mathbf{N}_1, \mathbf{S}) \le C_s(\widetilde{\mathbf{N}}_1, \mathbf{N}_1, \mathbf{S}) \tag{55}$$

$$= \frac{1}{2}\log\left|\frac{\mathbf{S}+\mathbf{N}_2}{\mathbf{B}^\star+\mathbf{N}_2}\right| - \frac{1}{2}\log\left|\frac{\mathbf{S}+\mathbf{N}_1}{\mathbf{B}^\star+\mathbf{N}_1}\right| \tag{56}$$

where the last equality follows from (54). This proved the desired converse result for $\mathbf{S} \succ 0$.

For the case when $\mathbf{S} \succeq 0$ but $|\mathbf{S}| = 0$, let

$$\theta = \mathsf{Rank}(\mathbf{S}) < t. \tag{57}$$

Following the same footsteps as in the proof of [14, Lemma 2], we can define an *equivalent* aligned MIMO Gaussian wiretap channel with θ transmit and receive antennas and a new covariance matrix power constraint that is strictly positive definite. Hence, we can convert the case when $\mathbf{S} \succeq 0$ but $|\mathbf{S}| = 0$ to the case when $\mathbf{S} \succ 0$ with the same secrecy capacity. This argument can be formally described as follows.

Since $\mathbf{S}$ is positive semidefinite, we can write

$$\mathbf{S} = \mathbf{Q_S}\mathbf{\Lambda_S}\mathbf{Q_S}^{\mathsf{T}} \tag{58}$$

where $\mathbf{Q_S}$ is an orthogonal matrix and

$$\mathbf{\Lambda_S} = \mathsf{Diag}(\underbrace{0, \dots, 0}_{t-\theta}, s_1, \dots, s_\theta) \tag{59}$$

is diagonal with $s_j > 0, j = 1, \dots, \theta$. For $k = 1, 2$, write

$$\mathbf{Q_S^{\mathsf{T}}}\mathbf{N}_k\mathbf{Q_S} = \begin{pmatrix} \mathbf{C}_k & \mathbf{D}_k \\ \mathbf{D}_k^{\mathsf{T}} & \mathbf{E}_k \end{pmatrix} \tag{60}$$

where $\mathbf{C}_k, \mathbf{D}_k$ and $\mathbf{E}_k$ are (sub)matrices of size $(t-\theta)\times(t-\theta)$, $(t-\theta)\times\theta$ and $\theta\times\theta$, respectively. Let

$$\mathbf{A}_k := \begin{pmatrix} \mathbf{I}_{t-\theta} & 0_{(t-\theta)\times\theta} \\ -\mathbf{D}_k^{\mathsf{T}}\mathbf{C}_k^{-1} & \mathbf{I}_\theta \end{pmatrix}, \quad k = 1, 2. \tag{61}$$

We now define an intermediate and equivalent channel by multiplying both sides of (38) by an *invertible* matrix $\mathbf{A}_k\mathbf{Q_S^{\mathsf{T}}}$

$$\mathbf{Y}'_k[m] = \mathbf{X}'[m] + \mathbf{Z}'_k[m], \quad k = 1, 2 \tag{62}$$

where

$$\mathbf{Y}'_k[m] = \mathbf{A}_k\mathbf{Q_S^{\mathsf{T}}}\mathbf{Y}_k[m] \tag{63}$$

$$\mathbf{X}'[m] = \mathbf{A}_k\mathbf{Q_S^{\mathsf{T}}}\mathbf{X}[m] \tag{64}$$

and

$$\mathbf{Z}'_k[m] = \mathbf{A}_k\mathbf{Q_S^{\mathsf{T}}}\mathbf{Z}_k[m]. \tag{65}$$

Then, the covariance matrix $\mathbf{N}'_k$ of the additive Gaussian noise vector $\mathbf{Z}'_k[m]$ is given by

$$\mathbf{N}'_k = \begin{pmatrix} \mathbf{C}_k & 0 \\ 0 & \mathbf{E}_k - \mathbf{D}_k^{\mathsf{T}}\mathbf{C}_k^{-1}\mathbf{D}_k \end{pmatrix} \tag{66}$$

and the matrix power constraint (2) becomes

$$\frac{1}{n}\sum_{m=1}^{n}\mathbf{X}'[m]\mathbf{X}'^{\mathsf{T}}[m] \preceq \mathbf{S}' \tag{67}$$

where

$$\begin{aligned} \mathbf{S}' &= \mathbf{A}_k\mathbf{Q_S^{\mathsf{T}}}\mathbf{S}\mathbf{Q_S}\mathbf{A}_k^{\mathsf{T}} \\ &= \mathbf{A}_k\mathbf{\Lambda_S}\mathbf{A}_k^{\mathsf{T}} \\ &= \mathbf{\Lambda_S}. \end{aligned} \tag{68}$$

Note from (68) that $\mathbf{S}'$ is diagonal with first $t-\theta$ diagonal elements equal to zero. Thus, the matrix constraint (67) requires that the first $t-\theta$ elements of $\mathbf{X}'[m]$ be zero. Moreover, from (66), the first $t-\theta$ and the rest of θ elements of $\mathbf{Z}'_k[m]$ are uncorrelated and hence must be independent as $\mathbf{Z}'_k[m]$ is Gaussian. Therefore, only the latter θ antennas transmit/receive information regarding message W. This allows us to define another *equivalent* aligned MIMO Gaussian broadcast channel with θ antennas at the transmitter and each of the receivers

$$\overline{\mathbf{Y}}_k[m] = \overline{\mathbf{X}}[m] + \overline{\mathbf{Z}}_k[m], \quad k = 1, 2 \tag{69}$$

where

$$\overline{\mathbf{Y}}_k[m] = \overline{\mathbf{A}}\mathbf{Y}'_k[m] \tag{70}$$

$$\overline{\mathbf{X}}[m] = \overline{\mathbf{A}}\mathbf{X}'[m] \tag{71}$$

$$\overline{\mathbf{Z}}_k[m] = \overline{\mathbf{A}}\mathbf{Z}'_k[m] \tag{72}$$

and $\overline{\mathbf{A}} = [0_{\theta\times(t-\theta)}\ \mathbf{I}_\theta]$. Now, the matrix power constraint (67) becomes

$$\frac{1}{n}\sum_{m=1}^{n}\overline{\mathbf{X}}[m]\overline{\mathbf{X}}^{\mathsf{T}}[m] \preceq \overline{\mathbf{S}} \tag{73}$$

where

$$\begin{aligned}\overline{\mathbf{S}} &= \overline{\mathbf{A}}\mathbf{S}'\overline{\mathbf{A}}^{\mathsf{T}}\\ &= \mathsf{Diag}(s_1,\ldots,s_\theta).\end{aligned} \tag{74}$$

Note that the matrix power constraint $\overline{\mathbf{S}}$ is *strictly* positive definite, so we can apply the previous result to the new wiretap channel (69). This completes the proof of the lemma. ∎

2) General MIMO Gaussian Wiretap Channel: For the general case, we may assume that the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are square but not necessarily invertible. If that is not the case, we can use singular value decomposition (SVD) to show that there is an equivalent channel which does have $t\times t$ square channel matrices. That is, we can find a new channel with square channel matrices which are derived from the original ones via matrix multiplications. The new channel is equivalent to the original one in preserving the secrecy capacity under the same power constraint.

Consider using SVD to write the channel matrices as follows:

$$\mathbf{H}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{V}_k^{\mathsf{T}}, \quad k=1,2 \tag{75}$$

where $\mathbf{U}_k$ and $\mathbf{V}_k$ are $t\times t$ orthogonal matrices, and $\mathbf{\Lambda}_k$ is diagonal. We now define a new MIMO Gaussian broadcast channel which has invertible channel matrices

$$\mathbf{Y}_k[m] = \overline{\mathbf{H}}_k\mathbf{X}[m] + \mathbf{Z}_k[m], \quad k=1,2 \tag{76}$$

where

$$\overline{\mathbf{H}}_k = \mathbf{U}_k(\mathbf{\Lambda}_k + \alpha\mathbf{I}_t)\mathbf{V}_k^t \tag{77}$$

for some $\alpha > 0$, and $\{\mathbf{Z}_k[m]\}_m$ is an i.i.d. additive vector Gaussian noise process with zero mean and identity covariance matrix. Note that the channel matrices $\overline{\mathbf{H}}_k$, $k=1,2$, are invertible. By Lemma 1, the secrecy capacity $C_s(\overline{\mathbf{H}}_2,\overline{\mathbf{H}}_1,\mathbf{S})$ of (38) (viewed as a MIMO Gaussian wiretap channel with receiver 2 as legitimate receiver and receiver 1 as eavesdropper) under the matrix power constraint (2) is given by

$$\begin{aligned}&C_s(\overline{\mathbf{H}}_2,\overline{\mathbf{H}}_1,\mathbf{S})\\ &= \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\frac{\mathbf{I}_t+\overline{\mathbf{H}}_2\mathbf{S}\overline{\mathbf{H}}_2^{\mathsf{T}}}{\mathbf{I}_t+\overline{\mathbf{H}}_2\mathbf{B}\overline{\mathbf{H}}_2^{\mathsf{T}}}\right| - \frac{1}{2}\log\left|\frac{\mathbf{I}_t+\overline{\mathbf{H}}_1\mathbf{S}\overline{\mathbf{H}}_1^{\mathsf{T}}}{\mathbf{I}_t+\overline{\mathbf{H}}_1\mathbf{B}\overline{\mathbf{H}}_1^{\mathsf{T}}}\right|\right).\end{aligned} \tag{78}$$

Finally, let $\alpha\downarrow 0$. We have $\overline{\mathbf{H}}_k\to\mathbf{H}_k$, $k=1,2$ and hence

$$\begin{aligned}&C_s(\overline{\mathbf{H}}_2,\overline{\mathbf{H}}_1,\mathbf{S})\\ &\to \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\frac{\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{S}\mathbf{H}_2^{\mathsf{T}}}{\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\mathsf{T}}}\right| - \frac{1}{2}\log\left|\frac{\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{S}\mathbf{H}_1^{\mathsf{T}}}{\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\mathsf{T}}}\right|\right).\end{aligned} \tag{79}$$

Moreover, by [9, Eqns. (45) and (46)]

$$C_s(\mathbf{H}_2,\mathbf{H}_1,\mathbf{S}) \le C_s(\overline{\mathbf{H}}_2,\overline{\mathbf{H}}_1,\mathbf{S}) + \mathcal{O}(\alpha) \tag{80}$$

where $\mathcal{O}(\alpha)\to 0$ in the limit as $\alpha\downarrow 0$. Thus, we have the desired converse result

$$\begin{aligned}&C_s(\mathbf{H}_2,\mathbf{H}_1,\mathbf{S})\\ &\le \max_{0\preceq\mathbf{B}\preceq\mathbf{S}}\left(\frac{1}{2}\log\left|\frac{\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{S}\mathbf{H}_2^{\mathsf{T}}}{\mathbf{I}_{r_2}+\mathbf{H}_2\mathbf{B}\mathbf{H}_2^{\mathsf{T}}}\right| - \frac{1}{2}\log\left|\frac{\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{S}\mathbf{H}_1^{\mathsf{T}}}{\mathbf{I}_{r_1}+\mathbf{H}_1\mathbf{B}\mathbf{H}_1^{\mathsf{T}}}\right|\right)\end{aligned} \tag{81}$$

by letting $\alpha\downarrow 0$ on the RHS of (80). This completes the proof of the theorem.

Appendix C
Proof of Theorem 3

In this Appendix, we prove Theorem 3. Without loss of generality, we may assume that the matrix power constraint $\mathbf{S}$ is strictly positive definite and the channel matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ are square but not necessarily invertible. We start with the following simple lemma.

Lemma 2: For any $t\times t$ matrices $\mathbf{B}$ and $\mathbf{H}$ such that $\mathbf{B}\succeq 0$, we have

$$|\mathbf{I}_t + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}| = |\mathbf{I}_t + \mathbf{H}^{\mathsf{T}}\mathbf{H}\mathbf{B}|. \tag{82}$$

In particular, if $\mathbf{B}=\mathbf{I}_t$, we have

$$|\mathbf{I}_t + \mathbf{H}\mathbf{H}^{\mathsf{T}}| = |\mathbf{I}_t + \mathbf{H}^{\mathsf{T}}\mathbf{H}|. \tag{83}$$

Proof: Note that if $\mathbf{H}$ is invertible, the equalities in (82) and (83) are trivial. Otherwise, consider using SVD to rewrite $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathsf{T}} \tag{84}$$

where $\mathbf{U}$ and $\mathbf{V}$ are $t\times t$ orthogonal matrices, and

$$\mathbf{\Lambda} = \mathsf{Diag}(\underbrace{0,\ldots,0}_{t-b},\lambda_1,\ldots,\lambda_b) \tag{85}$$

is diagonal with $\lambda_j > 0$, $j=1,\ldots,b$. Write

$$\mathbf{V}^{\mathsf{T}}\mathbf{B}\mathbf{V} = \begin{pmatrix}\mathbf{C_B} & \mathbf{D_B}\\ \mathbf{D_B^{\mathsf{T}}} & \mathbf{E_B}\end{pmatrix} \tag{86}$$

where $\mathbf{C_B}$, $\mathbf{D_B}$ and $\mathbf{E_B}$ are (sub)matrices of size $(t-b)\times(t-b)$, $(t-b)\times b$ and $b\times b$, respectively. Then

$$\begin{aligned}|\mathbf{I}_t + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathsf{T}}| &= |\mathbf{I}_t + \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathsf{T}}\mathbf{B}\mathbf{V}\mathbf{\Lambda}\mathbf{U}^{\mathsf{T}}|\\ &= |\mathbf{I}_t + \mathbf{\Lambda}\mathbf{V}^{\mathsf{T}}\mathbf{B}\mathbf{V}\mathbf{\Lambda}|\\ &= \left|\mathbf{I}_b + \overline{\mathbf{\Lambda}}\mathbf{E_B}\overline{\mathbf{\Lambda}}\right|\end{aligned} \tag{87}$$

where $\overline{\mathbf{\Lambda}} = \mathsf{Diag}(\lambda_1, \dots, \lambda_b)$. On the other hand

$$\begin{aligned}|\mathbf{I}_t + \mathbf{H}^\mathsf{T}\mathbf{H}\mathbf{B}| &= \left|\mathbf{I}_t + \mathbf{V}\mathbf{\Lambda}^2\mathbf{V}^\mathsf{T}\mathbf{B}\right| \\ &= \left|\mathbf{I}_t + \mathbf{\Lambda}^2\mathbf{V}^\mathsf{T}\mathbf{B}\mathbf{V}\right| \\ &= \left|\mathbf{I}_b + \overline{\mathbf{\Lambda}}^2\mathbf{E}_\mathbf{B}\right| \\ &= \left|\mathbf{I}_b + \overline{\mathbf{\Lambda}}\mathbf{E}_\mathbf{B}\overline{\mathbf{\Lambda}}\right| \end{aligned} \tag{88}$$

where the last equality follows from the fact that $\overline{\mathbf{\Lambda}}$ is invertible. Putting together (87) and (88) proves the equality in (82). This completes the proof of the lemma. ∎

We are now ready to prove Theorem 3, following the approach of [10]. Let

$$\mathbf{O}_k := \mathbf{H}_k^\mathsf{T}\mathbf{H}_k \quad k = 1, 2 \tag{89}$$

and let $\mathbf{\Phi}$ denote the generalized eigenvalue matrix of the pencil

$$\left(\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{O}_1\mathbf{S}^{\frac{1}{2}},\ \mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{O}_2\mathbf{S}^{\frac{1}{2}}\right) \tag{90}$$

such that

$$\mathbf{\Phi} = \begin{pmatrix} \overline{\mathbf{\Phi}}_1 & 0 \\ 0 & \overline{\mathbf{\Phi}}_2 \end{pmatrix} \tag{91}$$

where $\overline{\mathbf{\Phi}}_1 = \mathrm{Diag}\{\phi_1, \dots, \phi_\rho\}$ and $\overline{\mathbf{\Phi}}_2 = \mathrm{Diag}\{\phi_{\rho+1}, \dots, \phi_t\}$. Let $\mathbf{G}$ be the corresponding generalized eigenvector matrix such that

$$\begin{aligned} &\mathbf{G}^\mathsf{T}\left(\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{O}_1\mathbf{S}^{\frac{1}{2}}\right)\mathbf{G} = \mathbf{\Phi} \\ \text{and}\quad &\mathbf{G}^\mathsf{T}\left(\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{O}_2\mathbf{S}^{\frac{1}{2}}\right)\mathbf{G} = \mathbf{I}_t. \end{aligned} \tag{92}$$

Now define

$$\widetilde{\mathbf{O}} := \mathbf{S}^{-\frac{1}{2}}\left[\mathbf{G}^{-\mathsf{T}}\begin{pmatrix} \overline{\mathbf{\Phi}}_1 & 0 \\ 0 & \mathbf{I}_{t-\rho} \end{pmatrix}\mathbf{G}^{-1} - \mathbf{I}_t\right]\mathbf{S}^{-\frac{1}{2}}. \tag{93}$$

Since the generalized eigenvalues are ordered as

$$\phi_1 \geq \cdots \geq \phi_\rho > 1 \geq \phi_{\rho+1} \geq \cdots \geq \phi_t > 0 \tag{94}$$

we have

$$\begin{pmatrix} \overline{\mathbf{\Phi}}_1 & 0 \\ 0 & \mathbf{I}_{t-\rho} \end{pmatrix} \succeq \mathbf{\Phi} \tag{95}$$

$$\text{and} \quad \begin{pmatrix} \overline{\mathbf{\Phi}}_1 & 0 \\ 0 & \mathbf{I}_{t-\rho} \end{pmatrix} \succeq \mathbf{I}_t. \tag{96}$$

Hence, by (92) and (93)

$$\widetilde{\mathbf{O}} \succeq \mathbf{O}_1 \quad \text{and} \quad \widetilde{\mathbf{O}} \succeq \mathbf{O}_2. \tag{97}$$

It follows that [see (98)–(103), shown at the bottom of the page], where (98) follows from (83) in Lemma 2; (99) follows from the definition of $\mathbf{O}_1$ in (89); (100) follows from the fact that $\mathbf{O}_1 \preceq \widetilde{\mathbf{O}}$; (101) follows from the fact that $\mathbf{O}_2 \preceq \widetilde{\mathbf{O}}$ so the maximization occurs at $\mathbf{B} = \mathbf{S}$; and (102) follows (92) and the definition of $\widetilde{\mathbf{O}}$ in (93).

To prove the reverse inequality, let $\mathbf{G} = [\mathbf{G}_1\, \mathbf{G}_2]$ where $\mathbf{G}_1$ and $\mathbf{G}_2$ are (sub)matrices of size $t \times \rho$ and $t \times \rho$, respectively, and let

$$\mathbf{B}^\star := \mathbf{S}^{\frac{1}{2}}\mathbf{G}\begin{pmatrix} \left(\mathbf{G}_1^\mathsf{T}\mathbf{G}_1\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix}\mathbf{G}^\mathsf{T}\mathbf{S}^{\frac{1}{2}}. \tag{104}$$

Then, $\mathbf{B}^\star$ is positive semidefinite. Moreover, we may verify that $\mathbf{B}^\star \preceq \mathbf{S}$ as follows. Note that $\mathbf{G}$ is invertible, so it is enough to show that

$$\begin{pmatrix} (\mathbf{G}_1^\mathsf{T}\mathbf{G}_1)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \preceq \left(\mathbf{G}^\mathsf{T}\mathbf{G}\right)^{-1}. \tag{105}$$

Note that

$$\mathbf{G}^\mathsf{T}\mathbf{G} = \begin{pmatrix} \mathbf{G}_1^\mathsf{T}\mathbf{G}_1 & \mathbf{G}_1^\mathsf{T}\mathbf{G}_2 \\ \mathbf{G}_2^\mathsf{T}\mathbf{G}_1 & \mathbf{G}_2^\mathsf{T}\mathbf{G}_2 \end{pmatrix}. \tag{106}$$

$$\begin{aligned} C_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) &= \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{H}_1\mathbf{B}\mathbf{H}_1^\mathsf{T}\right| - \frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{H}_2\mathbf{B}\mathbf{H}_2^\mathsf{T}\right|\right) \\ &= \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{B}^{\frac{1}{2}}\mathbf{H}_1^\mathsf{T}\mathbf{H}_1\mathbf{B}^{\frac{1}{2}}\right| - \frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{B}^{\frac{1}{2}}\mathbf{H}_2^\mathsf{T}\mathbf{H}_2\mathbf{B}^{\frac{1}{2}}\right|\right) && (98) \\ &= \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{B}^{\frac{1}{2}}\mathbf{O}_1\mathbf{B}^{\frac{1}{2}}\right| - \frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{B}^{\frac{1}{2}}\mathbf{O}_2\mathbf{B}^{\frac{1}{2}}\right|\right) && (99) \\ &\leq \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}}\left(\frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{B}^{\frac{1}{2}}\widetilde{\mathbf{O}}\mathbf{B}^{\frac{1}{2}}\right| - \frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{B}^{\frac{1}{2}}\mathbf{O}_2\mathbf{B}^{\frac{1}{2}}\right|\right) && (100) \\ &= \frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\widetilde{\mathbf{O}}\mathbf{S}^{\frac{1}{2}}\right| - \frac{1}{2}\log\left|\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}}\mathbf{O}_2\mathbf{S}^{\frac{1}{2}}\right| && (101) \\ &= \frac{1}{2}\log\left|\overline{\mathbf{\Phi}}_1\right| && (102) \\ &= \frac{1}{2}\sum_{j=1}^{\rho}\log\phi_j && (103) \end{aligned}$$

Using block inversion, we may obtain (107), shown at the bottom of the page, where

$$\mathbf{E}_{\mathbf{G}} = \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_2 - \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2. \tag{108}$$

Since $\mathbf{G}^{\mathsf{T}}\mathbf{G}$ is positive definite, we have $\mathbf{E}_{\mathbf{G}} \succ 0$ and, hence, see (109)–(111), shown at the bottom of the page.

By (99)

$$\begin{aligned} C_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) &= \max_{0 \preceq \mathbf{B} \preceq \mathbf{S}} \left(\frac{1}{2} \log \left| \mathbf{I}_t + \mathbf{B}^{\frac{1}{2}} \mathbf{O}_1 \mathbf{B}^{\frac{1}{2}} \right| - \frac{1}{2} \log \left| \mathbf{I}_t + \mathbf{B}^{\frac{1}{2}} \mathbf{O}_2 \mathbf{B}^{\frac{1}{2}} \right| \right) \\ &\geq \frac{1}{2} \log \left| \mathbf{I}_t + \mathbf{B}^{\star\frac{1}{2}} \mathbf{O}_1 \mathbf{B}^{\star\frac{1}{2}} \right| - \frac{1}{2} \log \left| \mathbf{I}_t + \mathbf{B}^{\star\frac{1}{2}} \mathbf{O}_2 \mathbf{B}^{\star\frac{1}{2}} \right| \\ &= \frac{1}{2} \log \left| \mathbf{I}_t + \mathbf{B}^{\star} \mathbf{O}_1 \right| - \frac{1}{2} \log \left| \mathbf{I}_t + \mathbf{B}^{\star} \mathbf{O}_2 \right| \end{aligned} \tag{112}$$

where the last equality follows from (82). From (92), we have

$$\begin{aligned} \mathbf{O}_1 &= \mathbf{S}^{-\frac{1}{2}} \left(\mathbf{G}^{-\mathsf{T}} \mathbf{\Phi} \mathbf{G}^{-1} - \mathbf{I}_t \right) \mathbf{S}^{-\frac{1}{2}} \\ \text{and} \quad \mathbf{O}_2 &= \mathbf{S}^{-\frac{1}{2}} \left(\mathbf{G}^{-\mathsf{T}} \mathbf{G}^{-1} - \mathbf{I}_t \right) \mathbf{S}^{-\frac{1}{2}}. \end{aligned} \tag{113}$$

Hence, see (114)–(117), shown at the bottom of the page. giving

$$|\mathbf{I}_t + \mathbf{B}^{\star}\mathbf{O}_1| = \left|\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right|^{-1} \left|\overline{\mathbf{\Phi}}_1\right|. \tag{118}$$

Similarly, we may obtain

$$\begin{aligned} &\mathbf{B}^{\star}\mathbf{O}_2 \\ &= \mathbf{S}^{\frac{1}{2}}\mathbf{G} \begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1} - \mathbf{I}_\rho & -\left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2 \\ 0 & 0 \end{pmatrix} \mathbf{G}^{-1}\mathbf{S}^{-\frac{1}{2}} \end{aligned} \tag{119}$$

and

$$|\mathbf{I}_t + \mathbf{B}^{\star}\mathbf{O}_2| = \left|\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right|^{-1}. \tag{120}$$

Substituting (118) and (120) into (112), we may obtain

$$\begin{aligned} C_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) &\geq \frac{1}{2}\log\left|\overline{\mathbf{\Phi}}_1\right| \\ &= \frac{1}{2}\sum_{j=1}^{\rho}\log\phi_j. \end{aligned} \tag{121}$$

Putting together (103) and (121) establishes the desired equality

$$C_s(\mathbf{H}_1, \mathbf{H}_2, \mathbf{S}) = \frac{1}{2}\sum_{j=1}^{\rho}\log\phi_j. \tag{122}$$

This completes the proof of the theorem.

$$(\mathbf{G}^{\mathsf{T}}\mathbf{G})^{-1} = \begin{pmatrix} (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} + (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2\mathbf{E}_{\mathbf{G}}^{-1}\mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} & (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2\mathbf{E}_{\mathbf{G}}^{-1} \\ \mathbf{E}_{\mathbf{G}}^{-1}\mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} & \mathbf{E}_{\mathbf{G}}^{-1} \end{pmatrix} \tag{107}$$

$$\begin{aligned} (\mathbf{G}^{\mathsf{T}}\mathbf{G})^{-1} - \begin{pmatrix} (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2\mathbf{E}_{\mathbf{G}}^{-1}\mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} & (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2\mathbf{E}_{\mathbf{G}}^{-1} \\ \mathbf{E}_{\mathbf{G}}^{-1}\mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} & \mathbf{E}_{\mathbf{G}}^{-1} \end{pmatrix} && (109) \\ &= \begin{pmatrix} \mathbf{I}_\rho & (\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2 \\ 0 & \mathbf{I}_{t-\rho} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{E}_{\mathbf{G}}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I}_\rho & 0 \\ \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1)^{-1} & \mathbf{I}_{t-\rho} \end{pmatrix} && (110) \\ &\succeq 0 && (111) \end{aligned}$$

$$\begin{aligned} \mathbf{B}^{\star}\mathbf{O}_1 &= \mathbf{S}^{\frac{1}{2}}\mathbf{G} \begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{G}^{\mathsf{T}} \left(\mathbf{G}^{-\mathsf{T}}\mathbf{\Phi}\mathbf{G}^{-1} - \mathbf{I}_t \right) \mathbf{S}^{-\frac{1}{2}} && (114) \\ &= \mathbf{S}^{\frac{1}{2}}\mathbf{G} \left[\begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{\Phi} - \begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \mathbf{G}^{\mathsf{T}}\mathbf{G} \right] \mathbf{G}^{-1}\mathbf{S}^{-\frac{1}{2}} && (115) \\ &= \mathbf{S}^{\frac{1}{2}}\mathbf{G} \left[\begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \overline{\mathbf{\Phi}}_1 & 0 \\ 0 & \overline{\mathbf{\Phi}}_2 \end{pmatrix} - \begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1 & \mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2 \\ \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_1 & \mathbf{G}_2^{\mathsf{T}}\mathbf{G}_2 \end{pmatrix} \right] \mathbf{G}^{-1}\mathbf{S}^{-\frac{1}{2}} && (116) \\ &= \mathbf{S}^{\frac{1}{2}}\mathbf{G} \begin{pmatrix} \left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1}\overline{\mathbf{\Phi}}_1 - \mathbf{I}_\rho & -\left(\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_1\right)^{-1}\mathbf{G}_1^{\mathsf{T}}\mathbf{G}_2 \\ 0 & 0 \end{pmatrix} \mathbf{G}^{-1}\mathbf{S}^{-\frac{1}{2}} && (117) \end{aligned}$$

APPENDIX D
PROOF OF COROLLARY 2

Proof: By Theorem 1, we only need to show that the secrecy capacity

$$C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S}) = \frac{1}{2} \sum_{j=\rho+1}^{t} \log \frac{1}{\phi_j}. \tag{123}$$

Consider the pencil

$$\left(\mathbf{I}_t + \mathbf{S}^{\frac{1}{2}} \mathbf{H}_2^\mathsf{T} \mathbf{H}_2 \mathbf{S}^{\frac{1}{2}},\ \mathbf{I}_t + \mathbf{S}^{\frac{1}{2}} \mathbf{H}_1^\mathsf{T} \mathbf{H}_1 \mathbf{S}^{\frac{1}{2}}\right). \tag{124}$$

Note that the pencils (28) and (124) are generated by the same pair of semidefinite matrices but with different order. Therefore, the generalized eigenvalues of the pencil (124) are given by

$$0 < \frac{1}{\phi_1} \leq \cdots \leq \frac{1}{\phi_\rho} < 1 \leq \frac{1}{\phi_{\rho+1}} \leq \cdots \leq \frac{1}{\phi_t}. \tag{125}$$

Applying Theorem 3 for $C_s(\mathbf{H}_2, \mathbf{H}_1, \mathbf{S})$ and noting the fact that $\log \frac{1}{\phi_j} = 0$ when $\phi_j = 1$ completes the proof of the corollary. ■

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[3] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Information theoretic security," *Found. Trends Commun. Inf. Theory*, vol. 5, pp. 355–580, 2008.

[4] Z. Li, W. Trappe, and R. D. Yates, "Secret communication via multi-antenna transmission," presented at the 41st Annu. Conf. Information Sciences and Systems, Baltimore, MD, Mar. 2007.

[5] A. Khisti and G. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, pp. 3088–3104, Jul. 2010.

[6] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, pp. 4033–4039, Sep. 2009.

[7] A. Khisti and G. W. Wornell, "The MIMOME channel," in *Proc. 45th Annu. Allerton Conf. Comm., Contr., Computing*, Monticello, IL, Sep. 2007, pp. 625–632.

[8] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, Jul. 2008, pp. 524–528.

[9] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[10] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "A MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, Mar. 2009, Special Issue on Wireless Physical Security.

[11] H. D. Ly, T. Liu, and Y. Liang, "Multiple-input multiple-output Gaussian broadcast channels with common and confidential messages," *IEEE Trans. Inf. Theory*, to be published.

[12] R. Liu and H. V. Poor, "Secrecy capacity region of a multi-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235–1249, Mar. 2009.

[13] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493–2507, Jun. 2008.

[14] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[15] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Comm.*, vol. 7, pp. 2180–2189, Jun. 2008.

[16] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2004.

[17] G. Strang, *Linear Algebra and Its Applications*. Wellesley, MA: Wellesley-Cambridge, 1998.

Ruoheng Liu (S'02–M'07) received the B.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 1998, the M.S. degree in signal and information engineering from the Chinese Academy of Sciences, Beijing, in 2001, and the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, in 2007.

During his Ph.D. studies, he was with the Wireless Information Network Laboratory (WINLAB) at Rutgers University. From 2007 to 2009, he was a Postdoctoral Research Associate at the Department of Electrical Engineering, Princeton University, Princeton, NJ. Since February 2010, he has been with Alcatel-Lucent technologies, Murray Hill, NJ, where he is a wireless system engineer. His research interests are in the areas of wireless communications, coding theory, and information theory.

Tie Liu (S'99–M'06) received the B.S. (1998) and M.S. (2000) degrees in electrical engineering, from the Tsinghua University, Beijing, China, and the M.S. degree in mathematics (2004) and the Ph.D. degree in electrical and computer engineering (2006) from the University of Illinois at Urbana-Champaign, Urbana.

Since August 2006, he has been with Texas A&M University, College Station, where he is currently an Assistant Professor of electrical and computer engineering. His research interests are in the field of information theory, wireless communication, and statistical signal processing.

Dr. Liu is a recipient of the M. E. Van Valkenburg Graduate Research Award (2006) from the University of Illinois at Urbana-Champaign and the Faculty Early Career Development (CAREER) Award (2009) from the National Science Foundation.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. His research interests are in the areas of stochastic analysis, statistical signal processing, and information theory, and their applications in wireless networks and related fields. Among his publications in these areas are the recent books *Quickest Detection* (Cambridge University Press, 2009), coauthored with O. Hadjiliadis, and *Information Theoretic Security* (Now Publishers, 2009), coauthored with Y. Liang and S. Shamai.

Dr. Poor is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering of the U.K. He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, and in 2004–2007 as the Editor-in-Chief of these TRANSACTIONS. He was the recipient of the 2005 IEEE Education Medal. Recent recognition of his work includes the 2007 Technical Achievement Award of the IEEE Signal Processing Society, the 2008 Aaron D. Wyner Distinguished Service Award of the IEEE Information Theory Society, and the 2009 Edwin Howard Armstrong Achievement Award of the IEEE Communications Society.

Shlomo Shamai (Shitz) (S'80–M'82–SM'88–F'94) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1975, 1981, and 1986, respectively.

From 1975–1985, he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986, he has been with the Department of Electrical Engineering, Technion, where he is now the William Fondiller Professor of Telecommunications. His research interests encompass a wide spectrum of topics in information theory and statistical communications.

Dr. Shamai (Shitz) is the recipient of the 2011 Claude E. Shannon Award. He is the recipient of the 1999 van der Pol Gold Medal of the Union Radio Scientifique Internationale (URSI) and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, the 2003 and the 2004 joint IT/COM societies paper award, the 2007 IEEE Information Theory Society Paper Award, the 2009 European Commission FP7, Network of Excellence in Wireless COMmunications (NEWCOM++) Best Paper Award, and the 2010 Thomson Reuters Award for International Excellence in Scientific Research. He is also the recipient of the 1985 Alon Grant for distinguished young scientists and the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY and has also served on the Board of Governors of the Information Theory Society.